****

**By**

**Sunil Raj Thota  
ALY6030 - Data Warehousing and SQL**

**Class Name: Fall 2021 CPS Quarter**

**Class Number: 202215**

**CRN: 70526**

**Week 1 Assignment 1**

**thota.su@northeastern.edu**

**Introduction:**

In this Assignment, The report is divided into three sections. The first two sections cover how to install the MySQL Server and MySQL Workbench on my Windows machine. I have also shared each step with dedicated screenshots and comments where ever required. For this report, I have SQL as the language to interprets the results by creating a database schema and defining the table attributes with their data types. In the given list of datasets, I have chosen the “Airport” data. I have also gone beyond and analyzed 3 business questions. This dataset has 221 records with 8 variables.

**Analysis:**

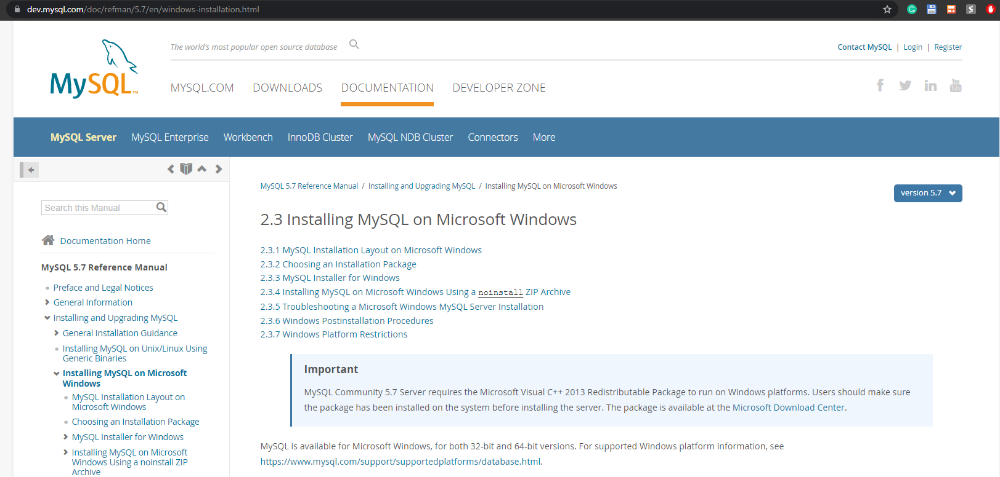
The Analysis part is divided into 3 sections;

**Part 1: Install MySQL Server**

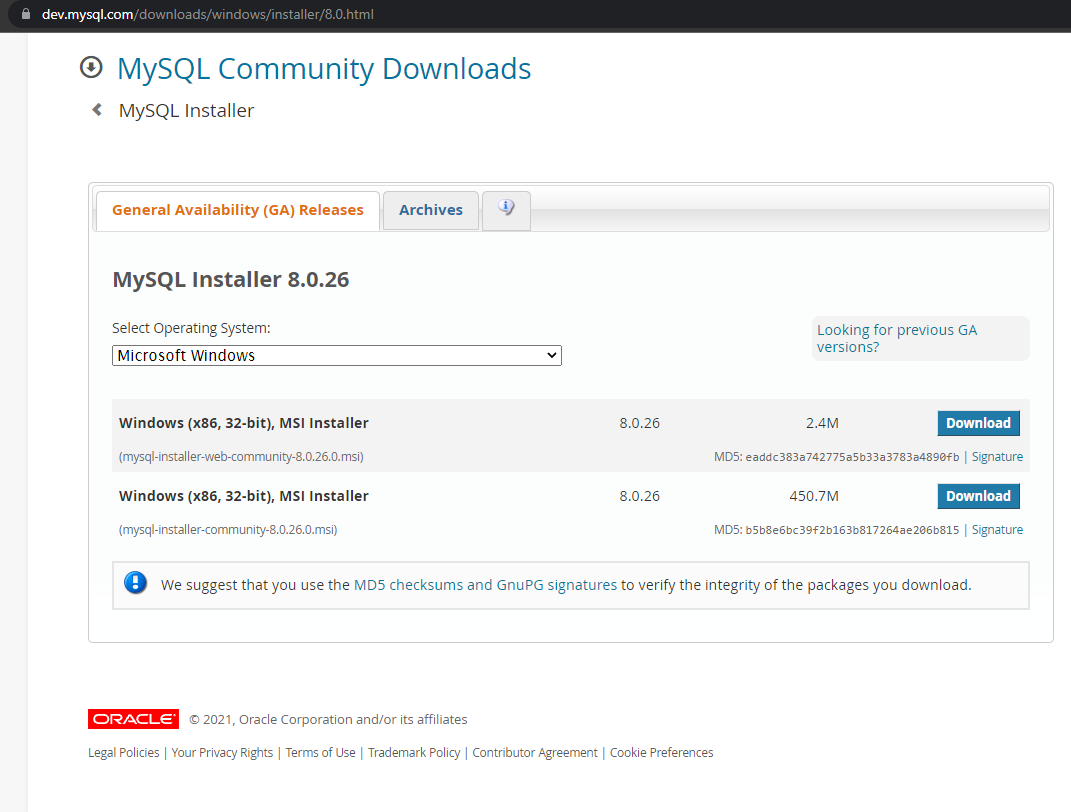
Step 1: Let’s click on the link provided to install the MySQL Server i.e., https://dev.mysql.com/doc/refman/5.7/en/installing.html



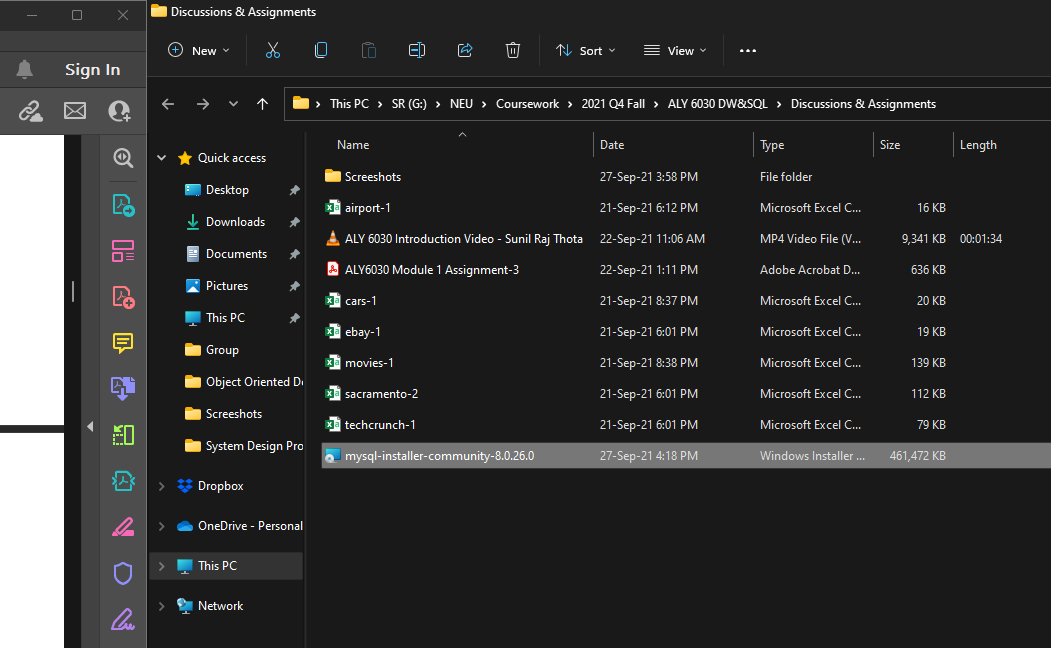
Step 2: Let’s click on the Installing MySQL on Microsoft Windows link.



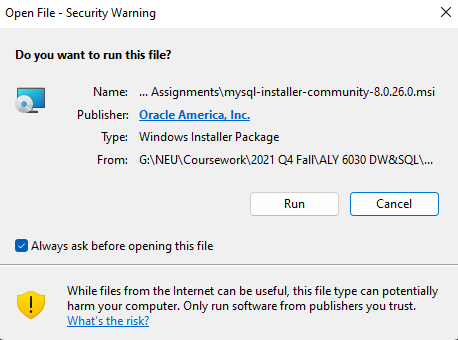
Step 3: Let’s download the necessary MSI Installers



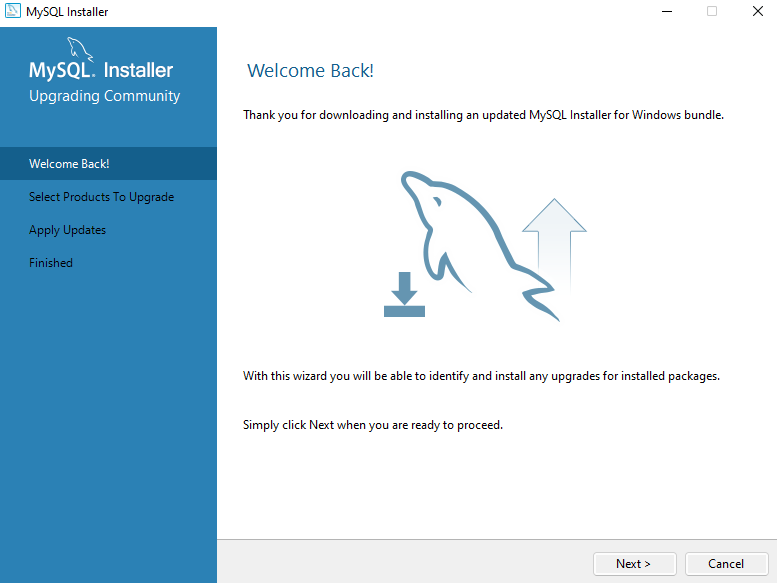
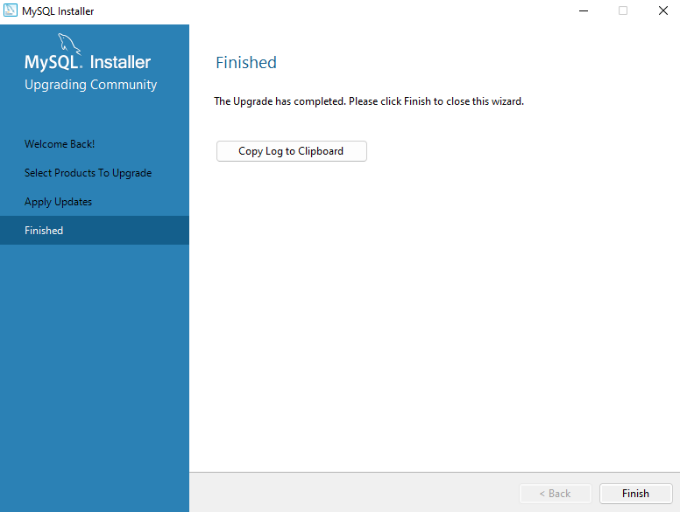
Step 4: In the below picture, we can observe the MySQL package installer



Step 5: Let’s open the file to follow the installation procedures.

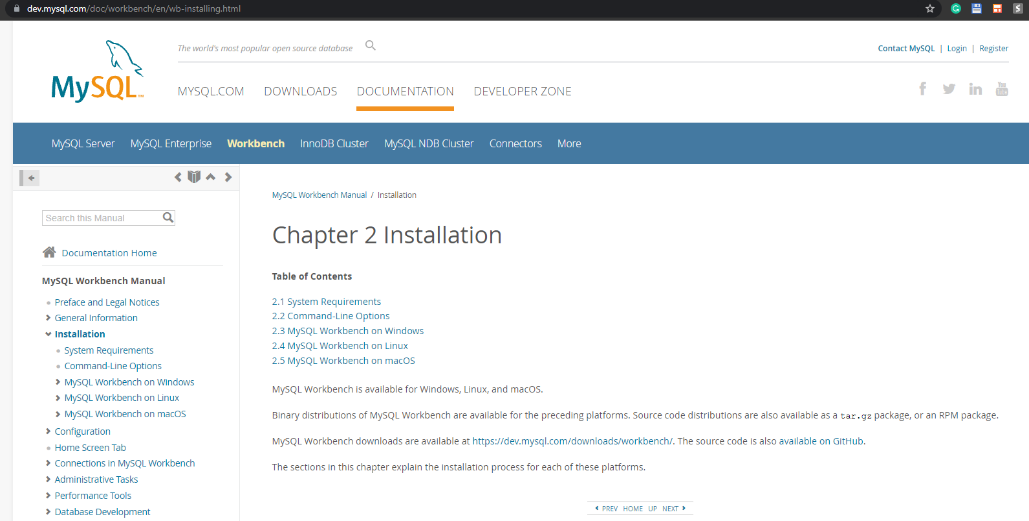


Step 6: In this let’s click on the Next Button till the software gets installed and then if it is installed already in our systems, we will be seeing something like this.

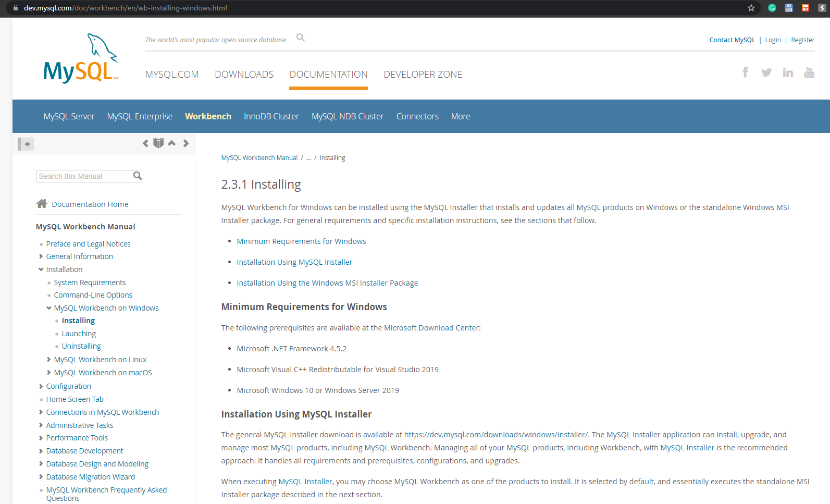
 

**Part 2: Install MySQL Workbench**

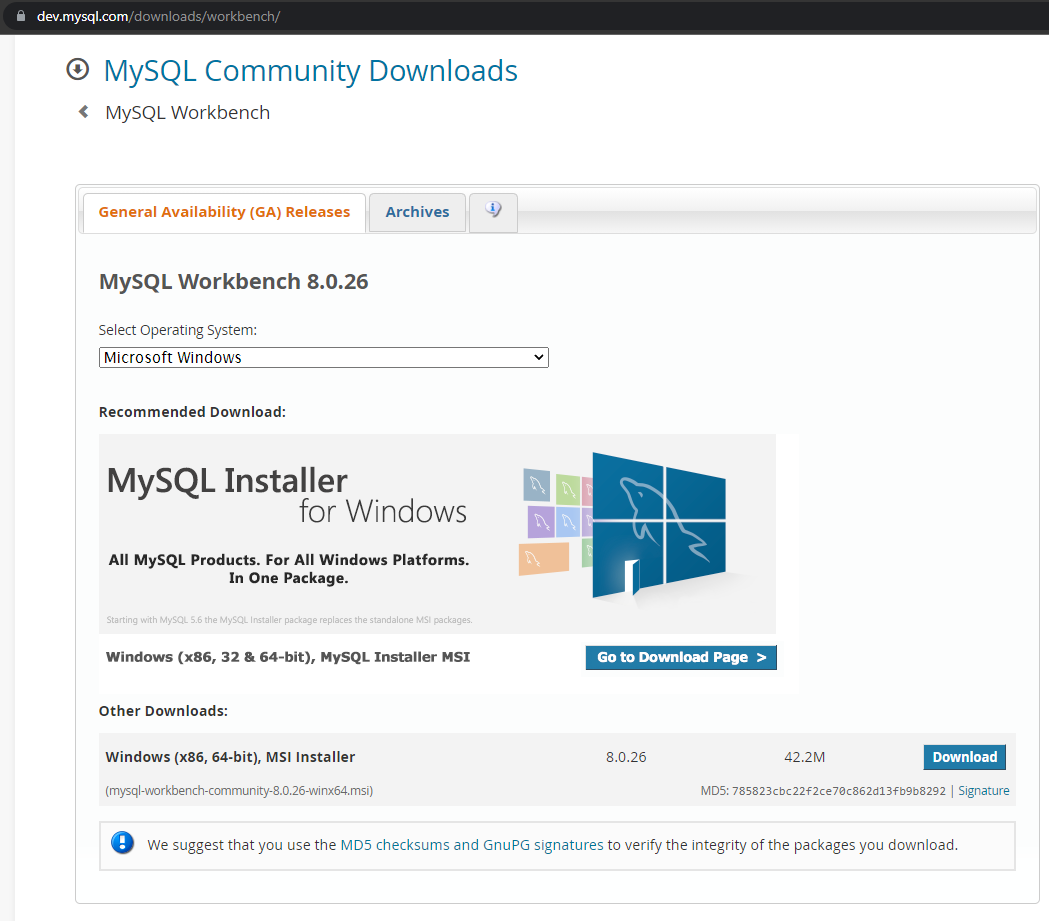
Step 1: Let’s click on the link provided to install the MySQL Server i.e., https://dev.mysql.com/doc/workbench/en/wb-installing.html



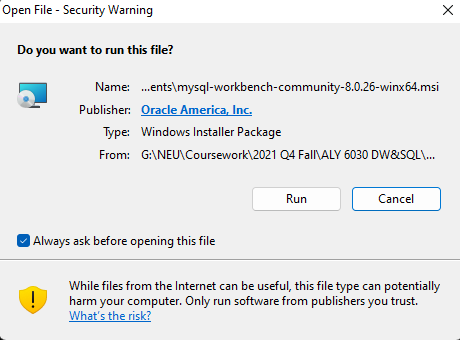
Step 2: Let’s follow the below process to install the MySQL Workbench CE



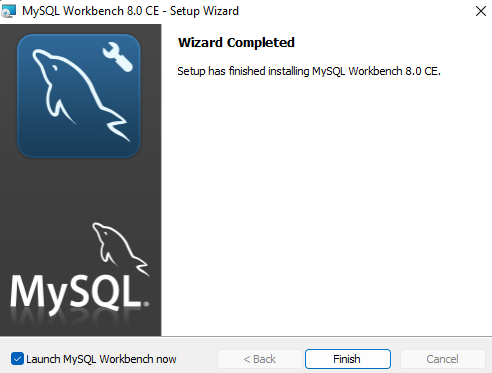
Step 3: Let’s download the MySQL Workbench 8.0.26



Step 4: Let’s open the file to follow the installation procedures.



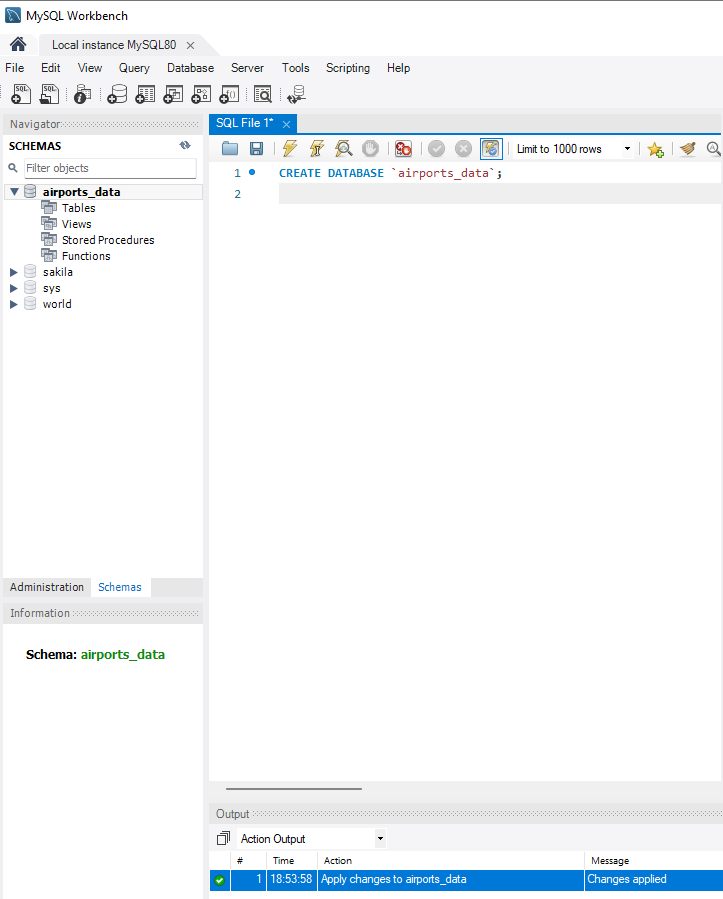
Step 5: We will see something like this once the installation is done.



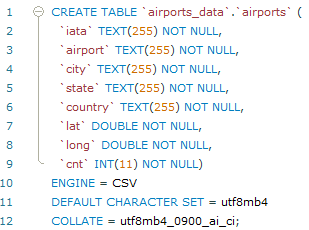
**Part 3: Data Modeling and SQL**

For the analysis part, I have chosen the “Airports” dataset

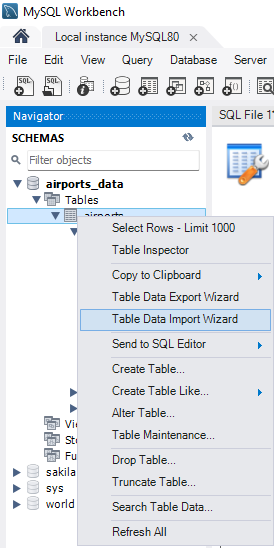
Step 1: After opening the MySQL Workbench and setting up the Server connection by giving a password. We will be seeing an IDE kind of workspace to do our analysis using SQL. Creating a database schema called “airports\_data”

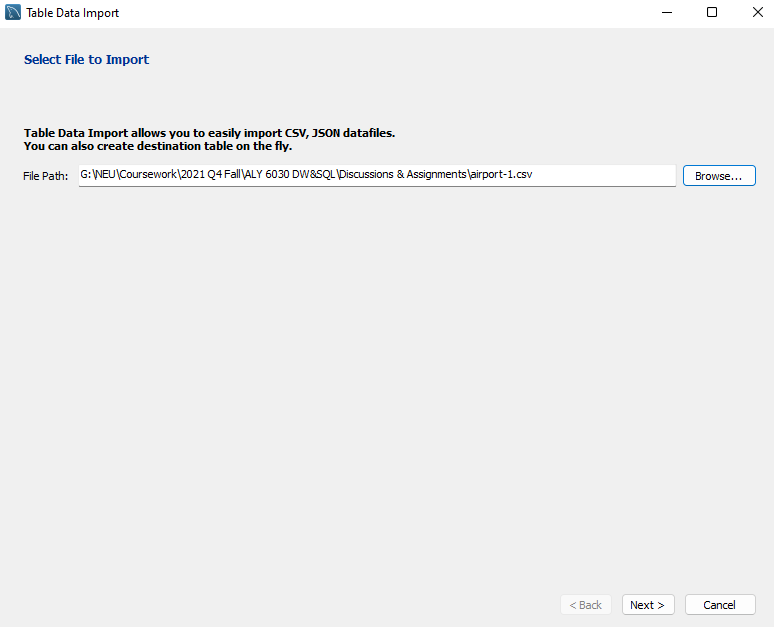


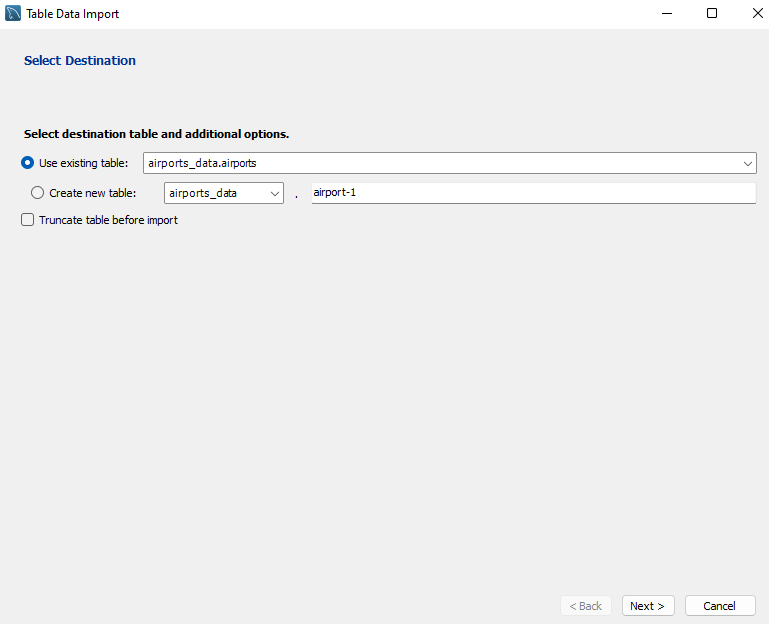
Step 2: Let’s create a table to store the Airports Data by defining their datatypes and the size of each attribute.

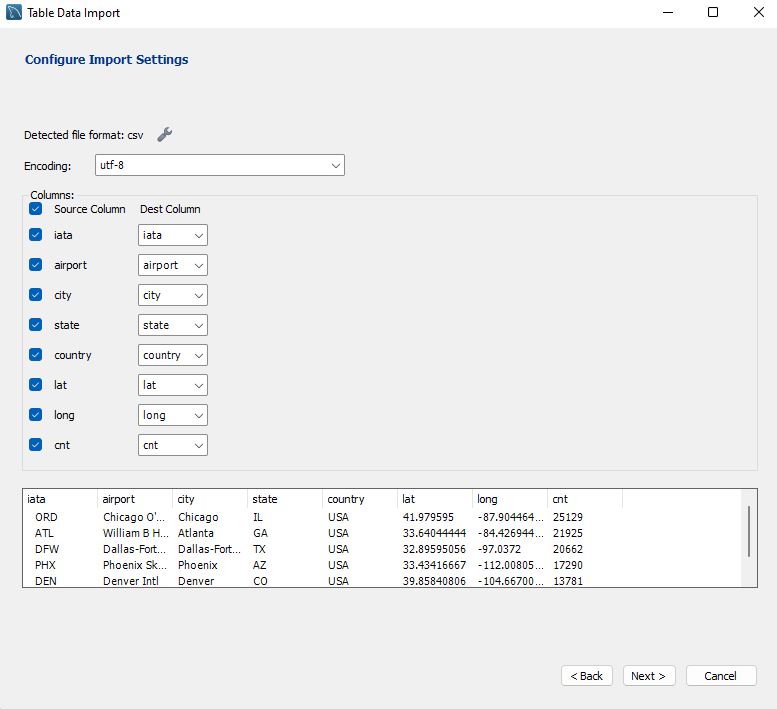


Step3: Next, step is to import the CSV dataset by using the “Table Data Import Wizard” option as shown below.

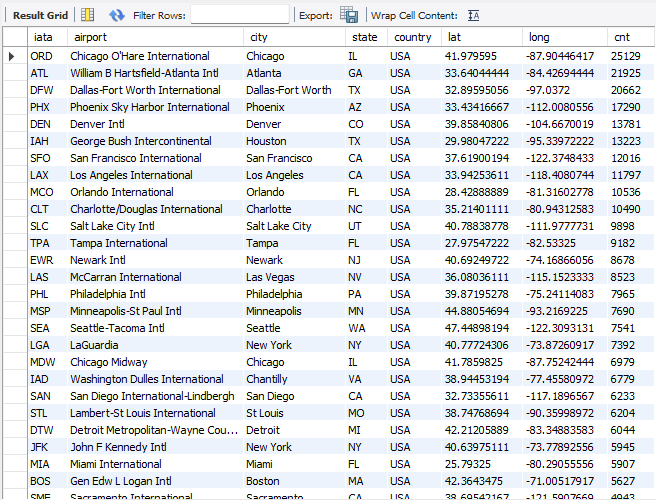




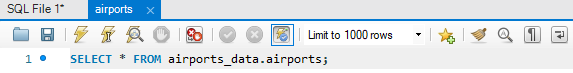




Step 4: This is how it looks once the dataset is imported and loaded.

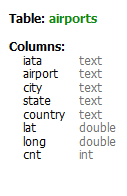


Step 5: This is to select all the records and display the result as shown in the above table.



**Questions:**

1. **What are the columns (attributes)?**



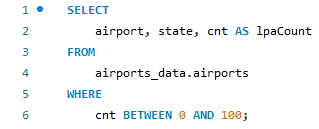
Here’s is a short description of each of the variables in the dataset:

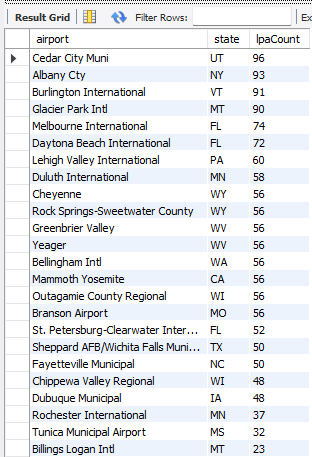
* **iata:** It is the airports abbreviation
* **airport:** Name of the airport
* **city:** city of the airport
* **state:** the state where the airport is located
* **country:** the country where the airport is located
* **lat:** latitude coordinate of the airport
* **long:** longitude coordinate of the airport
* **cnt:** number of flights that are taken care of by a particular airport authority

1. **What is the appropriate MySQL data type for each column?**

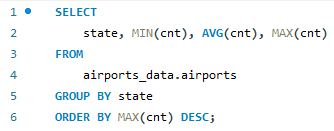
Every attribute has its datatype. TEXT() is given to those columns with Character Datatype are iata, airport, city, state, and country. Lat and long attributes were given with a datatype called DOUBLE. The last column cnt is given with the datatype called INT.

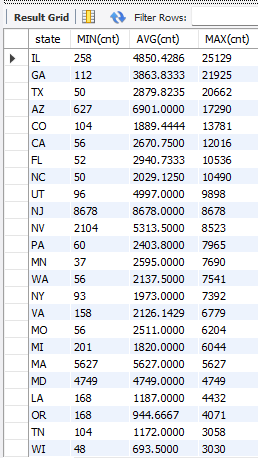
1. **Business Questions:**
2. **Which US airports have a count of low-performing airports that are less than or equal to 100?**



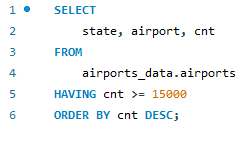


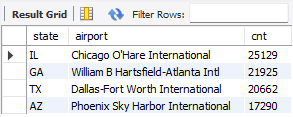
1. **What are the monthly minimum, average, and the maximum number of flights that are held by the USA states?**



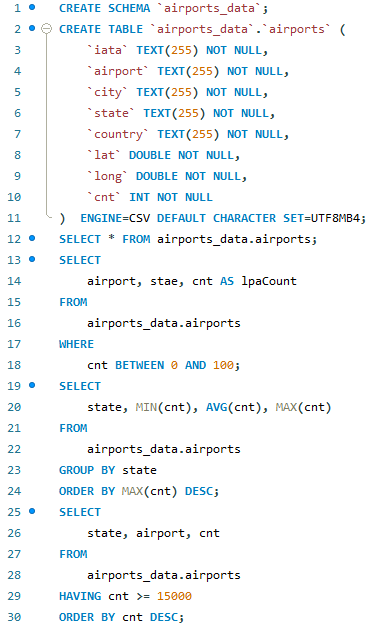


1. **Which US airports have a monthly count of more than 15,000 flights?**





Here is the complete SQL Code that is used to answer all the above questions



**Conclusion:**

In this assignment, we were able to create schemas, tables, and modified the data to obtain the appropriate outputs based on the queries and business questions that we have gone through. We have done the data analysis and modeling using the MySQL Workbench CE and SQL as the language. We also imported the CSV file and utilized the SQL Query editor which has some Intellisense. I have also shared all the necessary screenshots and output results with appropriate comments.

**Pros:** SQL Queries can be used for a variety of purposes from the most basic to the most complex. MySQL Server and MySQL Workbench software are widely utilized open-source community editions with the latest features. It can handle huge amounts of data. Easy to understand and the User Interface is simple to use.

**Cons:** MySQL installation has some issues in few Computers. Also, lack of stability in the software makes it tedious to troubleshoot